

## **StORe demonstrator – working specification**

### **Objectives**

1. Create permanent links between publications and underlying data
2. Create federations between institutional repositories, publishers and data archives
3. Within the federation only allow publication if data has been deposited or identified.
4. Make assignment of metadata simple
5. Have searchable metadata at collection level which is then inherited by items within the collection
6. Allow owners of collections to add individual items with at least titles as additional metadata.
7. Allow owners to determine which items are public and which are private.
8. Allow owners to form collaborations with colleagues and friends.
9. Allow owners to select items for publication
10. Allow owners to select items for deposit

### **Social science specific background**

1. Items will be stored under projects (the collections mentioned above)
2. Projects can be based on either a) primary or b) secondary analysis
  - Primary analysis creates new data
  - Secondary analysis is based on existing data and may or may not create new data
3. The source repository will be the UK Data Archive
4. The output repository will be the LSE Research Online
5. The institutional repository will be a test Fedora system at the University of Essex.
6. Authentication will be via Shibboleth

### **Technical Specification**

1. There should be 8 types of User
  - i StORe system administrator
  - ii Institutional Repository administrator
  - iii UKDA administrator
  - iv LSE administrator
  - v Authenticated user (collection owner)
  - vi Authenticated user (collection collaborator)
  - vii Authenticated user (collection guest)
  - viii Unauthenticated user
2. Unauthenticated user can only search
3. Authenticated user (collection guest) can search and read items in collections to which he has been added.
4. Authenticated user (collection collaborator) can search and add, edit and read items in collections to which he has been added.
5. Authenticated user (collection owner) can search, create collaborative collections and add, edit, delete, move, make public and read items in collections which owns.
6. LSE administrator is a special type of Authenticated user (collection owner) that has a 1 to 1 collaboration with ever Authenticated user (collection owner).

## **StORe demonstrator – working specification**

7. UKDA administrator is a special type of Authenticated user (collection owner) that has a 1 to 1 collaboration with ever Authenticated user (collection owner).
8. Institutional Repository administrator is the system administrator for a single institutional repository.
9. StORe system administrator is the system administrator for the central indexing and search interface.

N.B. In the demonstrator the search will be performed at the UKDA test institutional repository, hence 7, 8 and 9 above become the same person.

It is envisaged that when more repositories join the system these will be harvested and central Lucene search engine indexes will be made with a separate search interface.

10. There should be 3 types of item
  - i Private
  - ii Private until approved by authenticated user (collection owner), LSE administrator or UKDA administrator depending in which collection they are.
  - iii Public
11. Approval can mean one of three states as far as data is concerned.
  - i. Through an acquisition process, verified but rejected by source repository
  - ii. Through an acquisition process, verified and accepted by source repository
  - iii. Through an acquisition process, failed verification and rejected by source repository

N.B. in cases i and ii, item will become public and allow publications into an output repository to proceed.

In case iii the item remains private and publication into an output repository is not allowed.

## **Functional Specification**

1. A user coming to the StORe site can search collection level metadata and public items without logging in. They will also be able to view public items.
2. Any further activity is only allowed for users that have logged in.
3. The logging in process will require Shibboleth authorisation, although in early version of the system a login name and password will be used.
4. If logging in for the first time additional mandatory information will be requested Name, Organisation and Email address.

N.B. will probably need some acceptance that these can be displayed.

In final system an email will be sent with a link that will need to be clicked before the next stage can proceed.

5. If accepted an account is set up and 1 to 1 public collaborations with UKDA and LSE where user is an Authenticated user (collection guest) and the UKDA administrator and LSE administrator are respective Authenticated user (collection owners).
6. The authenticated user is now able to start creating collections or become collaborators or guests of existing collection if the owner of the collection so desires.

## **StORe demonstrator – working specification**

7. When starting a collection the user will be asked what type of analysis is involved in the project he is undertaking, ie primary, secondary creating primary or secondary.
8. If conducting a secondary analysis project he will be asked to supply at least 1 data Study Number (SN) of the source repository before the creation of the collection can proceed.

N.B. help will be given in identifying the SN and then a unique ID will be created which will be a URL linking to the catalogue page in the source repository for that data.

9. A title for the collection will then be requested.

N.B. the title and the SN can not be changed, although more SNs can be added.

10. Further collection level metadata items will be requested. Mandatory will be whether the collection is public or private, an abstract and some self assigned tags.

11. Items can then be added by the user or any collection collaborator.

N.B. in a public collection items added by a collection collaborator will be private until approved by the collection owner.

12. All items inherit the collection metadata and only an additional item title will be mandatory.

13. A collection owner can then try to move items, along with its associated metadata to his 1 to 1 collaboration with the output repository for peer-review and approval.

14. The outcome of this is dependant on the following:-

- i. If based on secondary analysis with no additional primary data being created the item is moved, but remains private until approved. (SN will already be attached)

- ii. If primary or producing additional data the corresponding collection name will be checked in the 1 to 1 collaboration with the source repository. If presented and verified (approved or rejected) then the item is moved, but remains private until approved. If the data item in the source repository is accepted then a SN will have been assigned and the unique URL identifier can be passed to the item waiting to be published. If the data item has been rejected then an acquisition number will be assigned to the item and no URL created.

N.B. if accepted it is envisaged that the LSE administrator would move the item to their source repository and leave a link to the item in the 1 to 1 collaboration. If rejected the item will be removed but the user is free to publish elsewhere or make the item public.

15. A collection owner can copy items, along with its associated metadata to his 1 to 1 collaboration with the source repository for verification and approval. He will be able to set an embargo period of time before the source repository will make it public (3mths, 6mths or 12mths?).

N.B. if accepted it is envisaged that the UKDA administrator would move the item to their output repository and leave a link to the item in the 1 to 1 collaboration. If verified but rejected then the item will remain in the 1 to 1 collaboration (self-archiving)

## **Search Functionality**

1. Simple search interface with Boolean operators.
2. Mechanism for searching on self assigned tags and other fields in the metadata eg users, organisation.
3. Owners can search for items added by collection collaborators.
4. Final version will have a mechanism by which the UKDA administrator and LSE administrator can trawl for new items added.